

1 VA

*Obrigatório

Endereço de e-mail *

Seu e-mail



Esta pergunta é obrigatória

Nome Completo *

Sua resposta



Esta pergunta é obrigatória

Um alquimista tem um vetor de string com nomes de substâncias (palavras mágicas) e um número após o nome que representa uma proporção em percentagem (%): `strg <- c("Voda 30", "bylinky 25", "ZEM 23", "zlAto 22")`. Em apenas uma linha de comando, usando as funções `tolower()` e `paste()`, escreva um código que deixe todas as palavras em letras minúsculas e com o sinal '%' após o número. Note que não pode haver espaço entre o número e a unidade de medida. Além disso, remova todos os espaços em branco do comando antes de submeter o mesmo. *

5 pontos

```
paste(tolower(strg), "%", sep="")
```



Qual(is) da(s) opção(ões) abaixo correspondem a regexp `"^([0-9][[:punct:]]+[a-z])"`? *

5 pontos

- ☒ 5!eAZ
- ☐ !6!rrrrDFGT
- ☐ R99Tf34!RR
- ☐ 123?aSd

Carregue o arquivo de texto que contém um trecho do romance "Gambler" de Fyodor Dostoyevsky. Após isso, submeta quantos caracteres tem nesse texto. Mas note que é importante importar o arquivo de texto como uma string de caractere única. Para isso use "read_file" do pacote "readr". *

5 pontos

2565

Submeta a regex que extrai todos os domínios que começam com http ou https do vetor a seguir: c ("www.dogman.com", "<http://rotterdam.com>", "<https://facebook.com>", "<httpx://sims.com>") *

5 pontos

```
grep(pattern="^https?:/",vetor,value=T)
```

Escreva uma função chamada 'ranqueamento' que gera o ranqueamento do dataset Forbes2000 de acordo com o valor de mercado (marketvalue). Esse ranqueamento pode ser tanto por categoria quanto geral. A função 'ranqueamento' recebe três argumentos: 'ranque', 'categoria' e 'opcao'. Se a 'opcao' 1 for escolhida, o ranqueamento geral deverá ser feito. Por outro lado, se a 'opcao' 2 for escolhida, o ranqueamento por categoria deverá ser realizado. Dado esses argumentos, a função retorna o nome da empresa, a categoria e o ranque. O protótipo dessa função é: `ranqueamento <- function(ranque, categoria=NULL, opcao){}`. Adicionalmente, se o raking ou a categoria não existir uma mensagem de erro deverá ser apresentada. De posse dessa função, responda as questões abaixo.



Exemplos de saídas da função.

```
> ranqueamento(ranque=1, opcao=1)
      name      category ranking
1 General Electric Conglomerates      1
> ranqueamento(ranque=1, categoria="Diversified financials", opcao=1)
      name      category ranking
1 General Electric Conglomerates      1
> ranqueamento(ranque=1, categoria="Software & services", opcao=2)
# A tibble: 1 x 3
# Groups:   category [1]
   name      category      ranking
  <fct>    <fct>        <int>
1 Microsoft Software & services      1
> ranqueamento(ranque=1, categoria="Comida e Serviço", opcao=2)
[1] "Categoria não existe!"
```

Qual o nome da empresa do seguinte código
ranqueamento(ranque=1234, opcao=1) ? *

5 pontos

- ☐ ExxonMobil
- ☐ Wal-Mart Stores
- ☒ OTP Bank
- ☐ Barclays
- ☐ NA
- ☐ "Categoria não existe!"



Qual a categoria da empresa do seguinte código ?
ranqueamento(ranque=198, opcao=1) *

5 pontos

- ☒ Media
- ☐ Wal-Mart Stores
- ☐ Barclays
- ☐ Diversified financials
- ☐ NA
- ☐ "Categoria não existe!"
- ☐ Banking

Qual o nome da empresa do seguinte código ?ranqueamento(ranque=45, categoria="Technology hardware & equipment", opcao=2) *

5 pontos

- ☐ Intel
- ☒ Chi Mei Optoelectronics
- ☐ Wal-Mart Stores
- ☐ Storage Technology
- ☐ NA
- ☐ "Categoria não existe!"
- ☐ Juniper Networks



Qual o nome da empresa do seguinte código ranqueamento(ranque=7, categoria="Bancos", opcao=2) ? * 5 pontos

- ☐ Intel
- ☐ Wachovia
- ☐ Wal-Mart Stores
- ☐ Barclays
- ☐ NA
- ☒ "Categoria não existe!"

Verifique o dataset diamonds que pode ser encontrado no pacote ggplot2. Você pode obter uma descrição dos dados digitando ?diamonds ou help(diamonds) no console. Note que para visualizar os dados você precisa instalar e carregar o pacote ggplot2. Também carregue o pacote dplyr. Após isso, responda as perguntas abaixo.

Qual o preço médio (price) dos diamantes cujo o corte (cut) é "Very Good" e o quilate (carat) é maior que 0.7. * 5 pontos

- ☒ 6512
- ☐ 5366
- ☐ 5997
- ☐ 7131
- ☐ 6952



Qual a cor (color) do diamante que possui o menor preço (price) com mais de 0.5 quilates (carat) ? *

5 pontos

- ☒ H
- ☐ D
- ☐ I
- ☐ J
- ☐ A
- ☐ B

Qual a proporção de Diamantes cujo o corte (cut) é "Premium" ? *

5 pontos

- ☐ 0.39953652
- ☐ 0.22398962
- ☐ 0.09095291
- ☐ 0.02984798
- ☒ 0.25567297

Instale e carregue o pacote AER e execute o comando `data("Fertility")` que carregará o dataset Fertility no seu RStudio. Esse dataset apresenta informações do censo americano de 1980 de mulheres casadas entre 21 e 35 anos com dois ou mais filhos. Resolva as questões abaixo. No entanto, primeiro verifique o dataset Fertility digitando `"?Fertility"` ou `"help(Fertility)"` no console.



Digite o código que seleciona as linhas de 35 a 50 das variáveis idade (age) e trabalho (work). Remova todos os espaços em branco, não altere o nome do dataset e use o dplyr. *

5 pontos

```
Fertility%>%select(age,work)%>%slice(35:50)
```

Qual a proporção de mulheres passaram a ter um terceiro filho (morekids) com mais de 30 semanas trabalhadas (work)? *

5 pontos

- ☐ 0.896809
- ☐ 0.998094
- ☐ 0.230398
- ☒ 0.103191
- ☐ 0.349505

Filtre o subconjunto de mulheres entre 22 e 24 anos e determine a proporção que tiveram menino como primogênito (gender1). Note que 22 e 24 deve ser incluído. *

5 pontos

- ☒ 0.5036608
- ☐ 0.5143607
- ☐ 0.5109608
- ☐ 0.512839
- ☐ 0.49839



Faça o download dos arquivos de dados Catfish.csv e Treatment.csv do Google classroom e importe-os para o R. Catfish.csv contém medições de peso (variáveis de março e abril) de duas espécies de bagres que são comumente usadas na aquicultura. Ambas as espécies foram criadas sob diferentes tratamentos de temperatura e alimentos, os quais são especificados em Treatment.csv. Obtenha uma visão geral da estrutura dos dados usando as funções `str()` e `head()`. Após isso, realize os seguintes tratamentos: (a) converta o dataframe do arquivo Catfish.csv em um formato longo, de modo que as medições de peso para março e abril sejam combinadas em uma coluna denominada Weight. (b) Una as colunas Genus e Species em uma variável chamada Species. O resultado deve ficar assim: *Silurus glanis*. (c) faça a combinação dos dados de Treatment.csv e Catfish.csv e salve o objeto resultante numa variável chamada Catfish_Treatment. A mesclagem deve ser feita através da coluna Tank. Em posse do novo Catfish_Treatment, responda as perguntas abaixo.

Qual o peso médio dos bagres do sexo masculino considerando o mês de Abril e o tratamento 2? 5 pontos *

- ☐ 71.1
- ☐ 72.5
- ☐ 24.2
- ☒ 72.1
- ☐ 23.8
- ☐ 48.1



Use a função "ifelse" para adicionar os elementos de uma nova coluna chamada de "AcimaMedia" ao dataset Catfish_Treatment. Essa coluna, que deve ser adicionada ao data frame, é uma coluna lógica, indicando "V" se o peso do bagre é maior que a média dos pesos. Caso contrário, atribua "F". Note que você deve tirar todos os espaços em branco e a coluna "AcimaMedia" deve ser atribuída ao data frame através do operador "\$". Além disso, as colunas do data frame devem ser acessadas pelo nome e usem aspas simples. *

5 pontos

```
Catfish_Treatment$AcimaMedia<-ifelse(Catfis
```

Com base na coluna que foi adicionada na questão anterior, responda:
Qual tank possui mais bagres com o peso acima da média? *

5 pontos

- ☐ Tank 1
- ☐ Tank 2
- ☐ Tank 3
- ☐ Tank 4
- ☐ NA
- ☒ Todos os Tanks possuem a mesma média.

Dada a tabela no link abaixo que representa uma competição entre Maria, Teresa, Francisca e Joaquina, determine quem foi a vencedora. 5 pontos

<https://docs.google.com/drawings/d/1-8biPvM3H8RaFFgLWttnnCuntSZ8lyKz8zvodemd5XY/edit?usp=sharing> *

- ☐ Maria
- ☐ Teresa
- ☐ Francisca
- ☒ Joaquina
- ☐ NA



Escreva uma função chamada 'casosCompleto' que retorna a quantidade de casos completos de um dado dataset . A função ' casosCompleto' recebe dois argumentos: 'df' e 'linhas'. Dado esses argumentos, a função retorna a quantidade de casos completos. O protótipo dessa função é:

casosCompleto<- function(df, linhas=NA){}. Se nenhuma linha for definida, então deve ser considerado todo o dataset. De posse dessa função, responda as questões abaixo.

Exemplo de saída da função.

```
> #####  
> set.seed(10)  
> df1 <- data.frame(matrix(data = sample(100,100,replace=TRUE), ncol = 10))  
> df1[df1>80] <- NA  
>  
> casosCompleto(df=df1)  
[1] "O numero de casos completos são: 3"  
> casosCompleto(df=df1, linhas = 7:10)  
[1] "O numero de casos completos são: 1"
```

Qual a quantidade de casos para o código a seguir?

5 pontos

casosCompleto(df=airquality) *

- ☐ 44
- ☐ 100
- ☐ 50
- ☒ 111
- ☐ NA
- ☐ 0



Qual a quantidade de casos completos para o código a seguir? Mas note 5 pontos
que "tb.csv" tem que está no diretório apontado pelo o RStudio. `aux<-
read.csv("tb.csv") casosCompleto(df=aux, linha=1348:4954).` *

- ☐ 456
- ☐ 392
- ☐ NA
- ☐ 241
- ☐ 0
- ☒ 177

Página 1 de 1

Enviar

Nunca envie senhas pelo Formulários Google.

Este formulário foi criado em UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO. [Denunciar abuso](#)

Google Formulários

